

1/4

**Figure 1A Nucleotide sequence of inserted environmental DNA (029cel)**  
**SEQ ID NO:1**

ATCAACACGC	TGGAAAGTAA	TTTCAAGGGT	AAGGCCATCG	GTTGCCGCCG	50
GGGTAGAAAT	GTGCGGTTGG	ATTTCTGTTGA	GCGGCGTCGC	CGGCGTTCCA	100
CCGAGGGCAT	AGCGCAGCAG	GTTGGCGATG	CCACCGGTGA	GGCCTTCGGG	150
GCCGCCTACG	ATGTTGTGCT	CAGCCGCCCA	TGCGATGTAG	CCGTCCGGCT	200
CGGGTTCGCT	CGCGGGGGTG	AAGAAGACAA	TGTCGTCTGAG	ATAAAGGTTG	250
CCGCTTCCGC	TCTCAACGCC	GCCGAGGTTG	AATTGGATTT	CGCAAATCT	300
CGTTAGGTCC	AGCACGGAAT	CGCCGACGAG	GTCGGCTATG	GGAATCTGAA	350
TGCGCCCAT	GGGTGTTGGTA	CGCGGAAGGG	ACACGTAGGG	ACCCACTTTG	400
TCATTGGGCG	AGACGAGCCG	GACAAAGATT	TGGTGCGCCG	CCTGCGAGGG	450
GCCTTGAGG	GCGAGAGAAA	GGTACGTGAG	GGCGCTGATG	TCGTGCGTGG	500
GACCGTCTCC	CCAGTTGTCTG	AGATTGAGCC	CAAATCCGGC	CCACCATCCG	550
GCGATAGTGT	AGCTCCAATG	GTAGTGACGC	TCACCTCGA	AGCCGCCGCT	600
GGAGAGTTCC	TGCAAGCCGT	CGCCCCAAAT	GCCCGTGATG	AGCGTTGCCT	650
CGTCACGTA	GATCACAAGT	TCGGCGGCGG	GTGCCGGGGG	AAGATCGCCT	700
TGAGTGATCA	CGAGAGTGGC	GGTGGCGCTG	CCTTCGTGAT	TAGGGTCGGT	750
AATGGTGCGG	ACGACCGTGT	AGCTACCGGG	CCCCACTGGC	GCATGGGTGG	800
AACCGTTGTA	GGTAAAGGAG	ACGTCAAGCC	CCACGGGATG	GGTCTCGGCA	850
AGAGCGGCCCT	TGGGGGTGCC	GTCGAAAACG	TGTTCCAAAT	TGGAGAGCGT	900
GATGGTGGCG	GGTGCCCTGA	GCACAGTCAC	AGAAACAGTG	GATTGCACGG	950
GATCGTGCGC	TGCCGTGTCT	GCAGGTGTGA	AGACCACGCT	GTAAAAACGG	1000
GTTCCGGCGG	ACGGTGCAAG	GCCGGACAGG	ACAAAGGCAA	AGTCGCCGGG	1050
GACGGCGGCT	ACTCCGCCGC	TCAGGCCGGC	CTCCGCAAGG	GTTTGCCCGA	1100
AGGTGATGGG	TGCGGCTGTG	GGCCACATCT	CCACAAGGCC	GGTGTCCCCC	1150
TCGTACGCA	CCGGCATGAG	GGCGGAGAGG	AGATGAATGT	AACTGGCTTG	1200
GTAATTGATG	TCGGGCTCGG	TGATTTCCCA	TGAGTTCTCC	GGCCAAAAAC	1250
CATTCCAATC	AAGGTAGGCT	TTTTGCACGG	GTTGGTCTCG	GATCGCCTGA	1300
ATGCTTCCGC	TGTATTTGGG	CATTGGGACC	CGCCCGAAAG	AAAACCAGGA	1350
GCGGGACCGT	AGAGTGAAGT	GAGGGCATTG	TCCCAGTCCG	GCCATCGCGG	1400
AACCAATGGT	GGTAGATTTT	ATTGGCTGCA	CGGTCAGCGC	CGCTGGCATA	1450
CATGTTGCTA	AGATAGACCA	TGCCCATTGG	GTTCACTCCG	TGGAGATAGT	1500
GCAGGTAGCC	CATCGCGGCA	TCGCGATGCG	CGGCCGCGTC	GGCGGGGTTG	1550
AGCCCAAGCC	TCCGTACCCC	CTCGAAGAAA	AAGCCAGCCT	GAGACTTTGT	1600
TTTGTTCGAG	CCCCACGTGT	AATCCTGATC	CTTCAGGTAG	GCGCGGTAGG	1650
CGTCGGTCTG	GTTATTCCAT	GCACCGAGAA	ACTCCCCACC	GTTTATAGAA	1700
GCCGCCATCC	GGTTGCGGAT	GTCGGCAGAG	ACGCTAGGCG	TCGCTCCCGG	1750
GAGGGTCGTG	TAGTGGGCGA	GAGCTTTTTG	TAGCTCACCT	TGAAAGGGGA	1800
AGAAATACCA	CCACTGCACG	GGCTCCATAT	CGAGATAGCG	CACATCGAAG	1850
AAATCGCGAT	AGACCGCACC	GCCCCTGCGC	TCGAAGAGCA	TGGCGGCGGC	1900
CATCACACGG	TTGGCTAGCG	TATCGTGGGC	ATTGCGCGAG	GGGCTCACGG	1950
AAGCAAATCC	GGTGTGTGTCG	AAAGGCACAT	GAGGATGGAC	CATGGTCCAA	2000
TTCCATGCGG	CGATGGCAGC	GGATTGAGG	GTGACGGCAT	AATCGCTCAT	2050
GCCTACGCTC	TCAAAGACAG	TCGCCCCGAG	GGCGAAAGCG	GCGGCAGCCA	2100
TGGCAGTGGC	CTCGGTCGAG	ACGGGGCCGT	AGTAACGCGG	ATGGGTGTCTG	2150
GTGCTCGGCG	GGCTGGCGCT	CTGGTGCCCC	GTCACGGAAA	CTTTCCCGAG	2200
AATAGCCCCG	CTCGGCTCCT	GCATGCGTAA	GAGCCAGTCC	ATTCCCCATT	2250
TGACTTCGTC	AAGCAGGTCTG	GGGACACCGT	TGCCGGATTTC	CGGGATGCCA	2300
AAATCATCGG	TAAAGACGTC	AGGCCGCCCT	TGATAGGCAA	GGAGCAGCTC	2350
CAGGATGACG	CGCCCCGTCC	ACTCGCTGTA	CTTGTTGAAA	TCGCCCCGCAT	2400

2/4

Figure 1B

CGAACCAACC	GCCGCTGAGA	TCGCGCTCCA	AGGAGGCATT	CCCCATATCC	2450
CAGATGGGGC	GGCTGGCGAC	GTCCTGCGGG	TGAGAAGCGG	CATCGGCCCA	2500
GTTCGCGTGG	GCGTAGGGCA	CCTCCTTGGC	AAACCCGGAG	CGCTGATAGA	2550
AGAACATGCG	CACGGCCTCG	CGCAGGACAA	CATCGTAAAC	ATCCGCGCCA	2600
ATGGCGAAAC	TATCGGAATG	AGTGTTGTTG	GCAGGATCGT	GGATGCGGTA	2650
GTGGCCGGGC	TCGGCAACTA	CCGTAAAATC	AAACCACCAC	ACGCGGTCTC	2700
CCGATTGAAT	ATGGATGGCG	CCGCCGTTCC	ACGGGACCGG	TGAGCCGGAG	2750
AAAACCACGA	CGCCATCGTT	CACGCGACGG	ACCTCCAGCG	TTGCGCCGGG	2800
GCTGTAGCTC	TCGGCGCTGT	TCCAGCCAAT	CTGCGGGTCG	GCGATCACCG	2850
CCACCTTGGT	GGCATCGGCG	GGGTAACCGA	ATTGGTCGAT	GCGGATTTTA	2900
TCGGTGTGGG	TGGAGGCGAC	GAGGGCGGAG	CTGCCCATGA	GCAGCAAGAA	2950
AAAGCCCGCT	GTCGGCCCGA	TACCAAAAAA	ACGAATAGGG	AGAGAAAAAT	3000
TCATAGCAGG	ATGTGGATAC	GGAAAGGGGG	AAAACGGTGC	AAAGACCCAA	3050
GCCCAACGCT	TGGCGAAAAC	TGGATGGTTG	GTTTATCAAG	AAAAGCGCTT	3100
TTGAGCCAAA	AGCTGCGGGC	AATCCTTATT	GCGTTTCACA	ATATTTTCAC	3150
ATCGTCGGCG	GCACGACTTT	TCGATGGGCG	ACTTGACAGC	GTATTCTCTC	3200
AGGCGCGAGG	CTGCAAACCT	TATGAAAAAA	GGCCCGCGCA	GCGATCTGTC	3250
CCCGGTCAAA	ATCCAGTCAA	GGTTTGTTCA	AGGGTTTGAG	GTCTGATAGA	3300
GGCACAGTCG	AGCCATCAGC	AGTCGCATTG	AGTAGGGTTG	TTGGAGAAAG	3350
TGTGCAAATG	ACCGCTGCCG	AAGGAACTGT	GGAGACAAAA	AGCATATTTT	3400
CCTCGCCAAG					3410

3/4

**Figure 2**  
**The nucleotide sequence of 029cel ORF**

ATGAATTTTT	CTCTCCCTAT	TCGTTTTTTT	GGTATCGGGC	CGACAGCGGG	50
CTTTTCTTTG	CTGCTCATGG	GCAGCTCCGC	CCTCGTCGCC	TCCACCCACA	100
CCGATAAAAT	CCGCATCGAC	CAATTCGGTT	ACCCCGCCGA	TGCCACCAAG	150
GTGGCGGTGA	TCGCCGACCC	GCAGATTGGC	TGGAACAGCG	CCGAGAGCTA	200
CAGCCCCGGC	GCAACGCTGG	AGGTCCGTCG	CGTGAACGAT	GGCGTCGTGG	250
TTTTCTCCGG	CTCACC GGTC	CCGTGGAACG	GCGGCGCCAT	CCATATTCAA	300
TCGGGAGACC	GCGTGTGGTG	GTTTGATTTT	ACGGTAGTTG	CCGAGCCCGG	350
CCACTACCGC	ATCCACGATC	CTGCCAACAA	CACTCATTC	GATAGTTTCG	400
CCATTGGCGC	GGATGTTTAC	GATGTTGTCC	TGCGCGAGGC	CGTGCGCATG	450
TTCTTCTATC	AGCGCTCCGG	GTTTGCCAAG	GAGGTGCCCT	ACGCCCACGC	500
GAAGTGGGCC	GATGCCGCTT	CTCACC CGCA	GGACGTCGCC	AGCCGCCCCA	550
TCTGGGATAT	GGGGAATGCC	TCCTTGGAGC	GCGATCTCAG	CGGCGGTTGG	600
TTCGATGCGG	GCGATTTC	CAAGTACAGC	GAGTGGACGG	GGCGCGTCAT	650
CCTGGAGCTG	CTCCTTGCC	ATCAAGGGCG	GCCTGACGTC	TTTACCGATG	700
ATTTTGGCAT	CCCGGAATCC	GGCAACGGTG	TCCCCGACCT	GCTTGACGAA	750
GTCAAATGGG	GAATGGACTG	GCTCTTACGC	ATGCAGGAGC	CGAGCGGGGC	800
TATTCTCGGG	AAAGTTTCCG	TGACGGGGCA	CCAGAGCGCC	AGCCCGCCGA	850
GCACCGACAC	CCATCCGCGT	TACTACGGCC	CCGTCTCGAC	CGAGGCCACT	900
GCCATGGCTG	CCGCCGCTTT	CGCCCTCGGG	GCGACTGTCT	TTGAGAGCGT	950
AGGCATGAGC	GATTATGCCG	TCACCCTCGA	ATCCGCTGCC	ATCGCCGCAT	1000
GGAATTGGAC	CATGGTCCAT	CCTCATGTGC	CTTTCGACAA	CACCGGATTT	1050
GCTTCCGTGA	GCCCCTCGCG	CAATGCCAC	GATACGCTAG	CCAACCGTGT	1100
GATGGCCGCC	GCCATGCTCT	TCGAGCGCAC	GGGCGGTGCG	GTCTATCGCG	1150
ATTTCTTCGA	TGTGCGCTAT	CTCGATATGG	AGCCCGTGCA	GTGGTGGTAT	1200
TTCTTCCCCT	TTCAAGGTGA	GCTACAAAAA	GCTCTCGCCC	ACTACACGAC	1250
CCTCCCGGGA	GCGACGCCTA	GCGTCTCTGC	CGACATCCGC	AACCGGATGG	1300
CGGCTTCTAT	AAACGGTG	GAGTTTCTCG	GTGCATGGAA	TAACCGAGACC	1350
GACGCCTACC	GCGCCTACCT	GAAGGATCAG	GATTACACGT	GGGGCTCGAA	1400
CAAAACAAAG	TCTCAGGCTG	GCTTTTTCTT	CGAGGGGGTA	CGGAGGCTTG	1450
GGCTCAACCC	CGCCGACGCG	GCCGCGCATC	GCGATGCCGC	GATGGGCTAC	1500
CTGCACTATC	TCCACGGAGT	GAACCCAATG	GGCATGGTCT	ATCTTAGCAA	1550
CATGTATGCC	AGCGGCGCTG	ACCGTG CAGC	CAATGAAATC	TACCACCATT	1600
GGTTCCGCGA	TGGCCGGACT	GGGACAATGC	CCTCACTTCA	CTCTACGGTC	1650
CCGCTCCTGG	TTTTCTTTTCG	GGCGGGTCCC	AATGCCCAAA	TACAGCGGAA	1700
GCATT CAGGC	GATCCGAGAC	CAACCCGTGC	AAAAAGCCTA	CCTTGA	1746

4/4

**Figure 3****The translated 029cel protein composed of 581 amino acids**

MNFSLPIRFF	GIGPTAGFFL	LLMGSSALVA	STHTDKIRID	QFGYPADATK	50
VAVIADPQIG	WNSAESYSPG	ATLEVRRVND	GVVVFSGSPV	PWNGGAIHIQ	100
SGDRVWWFDF	TVVAEPGHYR	IHDPANNTHS	DSFAIGADVY	DVVLREAVRM	150
FFYQRSGEFAK	EVPIAHANWA	DAASHPQDVA	SRPIWDMGNA	SLERDLSGGW	200
FDAGDFNKYS	EWTEGRVILEL	LLAYQGRPDV	FTDDFGIPES	GNGVPDLLDE	250
VKWGMDWLLR	MQEPSGAILG	KVSVTGHQSA	SPPSTDTHPR	YYGPVSTEAT	300
AMAAAAFALG	ATVFESVGMS	DYAVTLESAA	IAAWNWTMVH	PHVFPDNTGF	350
ASVSPSRNAH	DTLANRVMAA	AMLFERTGGA	VYRDFFDVRY	LDMEPVQWWY	400
FFPFQGELQK	ALAHYTTLPG	ATPSVSADIR	NRMAASINGG	EFLGAWNNQT	450
DAYRAYLKDQ	DYTWGSNKTG	SQAGFFFEQV	RRLGLNPADA	AAHRDAAMGY	500
LHYLHGVNPM	GMVYLSNMYA	SGADRAANEI	YHHWFRDGRT	GTMPSLHSTV	550
PLLVFFRAGP	NAQIQRKHSG	DPRPTRAQSL	P		581